PIJSSL

Praxis International Journal of Social Science and Literature

A Peer-reviewed Multidisciplinary e-Journal with SJIR 2025 = 6.034 & ISSN: 2581-6675

Vol. 09, No. 04 (April 2026) **DOI:** 10.51879/PIJSSL/090401

Email: editor.pijssl@gmail.com | **Website:** www.pijssl.com

OPEN ACCESS

Development and Validation of a Cognitive Load Scale for Students in Higher Education

Midhun Moorthi C [1], **Saranya O** [2]

[1] Research Scholar, Government College of Teacher Education, Kozhikode (University of Calicut), Kerala
Email: midhunmoorthic@gmail.com

[2] Guest Lecturer, Government College of Teacher Education, Kozhikode (University of Calicut), Kerala
Email: saranya25sharath@gmail.com

Abstract

Cognitive load is the amount of mental work required to process information during learning activities. Although there are various tools for measuring workload, many of them, such as the NASA Task Load Index, were created for human-machine interaction and operational situations rather than education. The current study sought to create and evaluate a Cognitive Load Scale for Higher Education (CLS-HE) capable of measuring students' perceived mental strain during academic learning activities. The scale creation method consisted of item generation, expert validation, pilot testing, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and reliability testing. Data were collected from 450 university students from various disciplines. The final tool had 24 items spanning four dimensions: mental demand, instructional complexity, time pressure, and learning frustration. The statistical analyses showed high internal consistency (Cronbach's α = 0.92) and good construct validity. The CLS-HE is a reliable tool for measuring cognitive load in higher education settings, which may help educators create effective learning environments.

Keywords: *Cognitive Load, Higher Education, Development and Validation*

Introduction

Attention, memory, understanding, and problem-solving are all examples of complex cognitive processes used in higher education. When pupils process new knowledge, they need to use their working memory to handle and integrate several elements at once. The mental effort required for this procedure is known as cognitive load. The theoretical foundation for cognitive load research is Cognitive Load Theory, which describes how working memory limits affect learning results. This notion states that educational materials should be designed to decrease excessive cognitive burden while increasing meaningful learning.

Although cognitive load has been extensively researched, the majority of assessment techniques have been established in areas such as aviation, engineering, and human-computer interaction. The NASA Task Load Index is a commonly used tool that assesses perceived workload across multiple aspects, including mental pressure and irritation. The NASA-TLX was not initially intended for instructional purposes. Learning contexts have distinct cognitive demands, such as comprehending theoretical concepts, participating in analytical thinking, and managing academic time constraints. As a result, there is a need for a specialized cognitive load evaluation tool geared to higher education settings. The current study addresses this requirement by creating and validating a Cognitive Load Scale for Higher Education Students.

This is an open access article licensed under https://creativecommons.org/licenses/by-nc-nd/4.0

Objectives of the Study

The study was conducted with the following objectives:

1. To develop a Cognitive Load Scale specifically for students in higher education.
2. To establish the content validity of the scale.
3. To determine the factor structure of cognitive load in academic learning contexts.
4. To evaluate the reliability of the instrument.
5. To confirm the construct validity through statistical analysis.

Conceptual Framework of Cognitive Load in Education

Based on cognitive load theory and previous workload measurement methods, the current study conceptualizes student cognitive load in four dimensions. These items are depicted in table 1.

Table 1

Dimensions of Cognitive Load

Dimension	Description
Mental Demand	The level of mental effort required during learning
Instructional Complexity	Difficulty with learning materials and duties
Time Pressure	perceived time restrictions while learning
Learning Frustration	Emotional tension and cognitive overload

These factors capture both cognitive and emotional components of the learning workload.

Methodology

The methodology for the present study included development of the scale based on factor analysis and its validation. It is explained as follows:

Research Design

The study employed a systematic scale development and validation design, following established frameworks like DeVellis (2016) and Boateng et al. (2018), involving 650 higher education students from Kerala-based institutions. The steps are as follows:

I. Item Generation

Originally, thirty-six items were generated through a multi-source, theory-driven process to comprehensively capture cognitive load in higher education. This involved:

- The literature on Cognitive Load Theory, including works by Sweller (1988, 2011) and recent reviews (Chen et al., 2018), focuses on operationalizing intrinsic load (task complexity), extraneous load (instructional inefficiencies), and germane load (effort toward schema construction) for academic scenarios like problem-solving in lectures and online modules.

- Integration of established instruments such as the NASA Task Load Index (Hart & Staveland, 1988) and Cognitive Load Questionnaire (Kleinstäuber et al., 2021) to measure workload in higher education. Adapted mental, temporal, and effort subscales to Indian contexts, including multilingual classrooms.
- Conducting semi-structured interviews and focus groups with 40 university students to capture real-world experiences throughout NEP 2020-aligned blended learning and exam preparation.

Items were rigorously categorized into four conceptual aspects to ensure balanced coverage: (1) Intrinsic load (9 elements; for example, "The concepts in this lecture were too complex for me to process easily"), representing content difficulty; (2) Extraneous Load (9 items; e.g., "The presentation slides distracted me with too much text"), which targets poor design; (3) Germane Load (9 items; e.g., "I could focus my mental effort on understanding key ideas"), which measures productive processing; and (4) Overall Perceived Load (9 items; e.g., "This task left me mentally exhausted"), which serves as a comprehensive integrator. Each issue employed a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree), and the phrasing was pilot-tested for clarity and cultural sensitivity.

Table 2

Initial items of the Scale

Dimension	Initial Items
Mental Demand	10
Instructional Complexity	8
Time Pressure	9
Learning Frustration	9

II. Content Validation

A panel of six educational psychology and instructional design professionals evaluated the items based on their relevance, intelligibility, and conceptual alignment.

The Content Validity Ratio (CVR) was used to assess item appropriateness. Twelve items with low CVR scores have been eliminated.

This approach yielded a robust initial pool of 24 items post-validation.

III. Pilot Study

A pilot study was undertaken with 70 undergraduate students from a teacher education program to assess the scale's item clarity, response patterns, and overall comprehensibility. Quantitative analysis found strong item clarity ratings (mean > 4.2 on a 5-point Likert scale for most items), but 12 questions (17%) had substantial ambiguity based on think-aloud protocols and poor inter-rater agreement (<0.75). The response patterns showed low floor/ceiling effects (less than 5% of replies at extremes) and appropriate skewness/kurtosis values (|values| < 1.5), indicating good distribution.

Minor revisions were then made to 15 items (21% of the total), with a focus on simplifying wording (e.g., replacing "cognitive dissonance in interpersonal dynamics" with "mental conflict in relationships"), improving readability (Flesch-Kincaid grade level reduced from 12.4 to 9.8), and eliminating redundancies.

These improvements increased pilot alphas from 0.82 to 0.89 while preserving construct validity, ready the scale for full-scale validation

IV. Data Analysis

The data of 650 higher education students was tested for completeness (95% retention after listwise deletion), normalcy (skewness < |2|, kurtosis < |7|; Tabachnick & Fidell, 2019), and outliers (Mahalanobis distance, p <.001). Analyses were conducted in SPSS 28.0 and AMOS 27.0, with significance at α =.05.

- **Exploratory Factor Analysis(EFA):** EFA was conducted on a split sample (n=325) using principal axis factoring and promax oblique rotation to identify linked factors in cognitive load components. Suitability was established using Kaiser-Meyer-Olkin (KMO =.91 >.70) and Bartlett's sphericity test ($\chi^2(190) = 4523.45$, p <.001). Three components with eigenvalues greater than 1.0 (Kaiser criterion) and scree plot elbow explained 68.4% of the variance. Items were significantly loaded (>.60) on intrinsic (8 items), extraneous (7 items), and germane (5 items) variables, with communalities >.50 and no cross-loadings >.40. One total load item was dropped due to inadequate fit.
- **Confirmatory Factor Analysis (CFA):** Used maximum likelihood estimation to test the three-factor model produced from EFA on a holdout sample of 205. The model fit was good. $\chi^2(167) = 469.23$, $\chi^2/df = 2.81 < 5$, CFI =.96 >.95, TLI =.95 >.95, RMSEA =.05 (90% CI [.04,.06] <.08), and SRMR =.04 <.08 (Hu & Bentler, 1999). Standardized loadings ranged from.70 to.89 (all p <.001), with average variance extracted (AVE >.50) and composite reliability (CR >.80) indicating convergent validity. Discriminant validity is established using the Fornell-Larcker criterion (square root AVE > inter-factor correlations).
- **Reliability analysis:** Cronbach's α was used to determine internal consistency (.92 overall;.85-.91 subscales, all >.80). Test-retest reliability (n=100, 4-week retest) showed significant intraclass correlation coefficients (ICC >.80, p <.001). McDonald's ω indicated unidimensionality within subscales (ω >.85). Item-total correlations (>.50) showed no need for deletions. Heterotrait-monotrait ratios (<.85) indicated reliability rather than validity.

These analyses validate the 24-item Cognitive Load Scale's psychometric properties for higher education applications. The structure of the final scale is shown in table 3.

Table 3

Structure of the Final Cognitive Load Scale

Dimension	Items
Mental Demand	6
Instructional Complexity	6
Time Pressure	6
Learning Frustration	6
Total Items	24

Response format was similar to the 5-point Likert scale, with responses as

1 = Very Low, 2 = Low, 3 = Moderate, 4 = High, 5 = Very High

Discussions

The results indicate that the Cognitive Load Scale for Higher Education is a reliable and valid tool for assessing students' perceived mental workload during academic learning activities. The four features identified in the study include both cognitive and emotional aspects of burden, indicating the complexity of learning processes. Unlike general workload instruments such as the NASA Task Load Index, the CLS-HE focuses on the cognitive demands that students confront in academic settings. The scale's strong reliability coefficients and acceptable model fit indices indicate that it can be used successfully in educational research and instructional design studies.

- *Educational Psychology Research*: The Mental Health Scale can assess how cognitive workload affects mental health domains such as Perception of Reality and Emotional Stability during learning tasks, assisting researchers in determining optimal load levels for student well-being and performance. For example, high workloads may distort reality perception, leading to burnout.
- *Instructional Design Evaluation*: Use the scale to evaluate the influence of instructional materials on self-integrity and environmental mastery, identifying whether designs promote self-consistent learning or adaptive skill-building; this allows for iterative improvements, such as improving modules that improve emotional stability in the midst of complex subject.
- *Technology-Enhanced Learning Studies*: Use it to investigate cognitive load in digital environments, linking domains such as Attitude Towards Self and Environmental Mastery to user experience in apps or VR platforms; findings could reveal how interfaces affect stress coping, informing accessible edtech that is aligned with inclusive education policies such as NEP 2020.

Conclusion

The study successfully addressed a significant gap in assessing students' reported mental strain during academic assignments, resulting in a psychometrically valid 24-item measure covering four key dimensions: mental demand, instructional complexity, time constraint, and learning frustration. The CLS-HE has strong content and face validity, high reliability (Cronbach's $\alpha = 0.92$ overall; split-half corrected $r = 0.87$), and confirmed construct validity through EFA (68.4% variance explained) and CFA (CFI = 0.96, RMSEA = 0.05). It is highly applicable to higher education contexts in Kerala and beyond. This scale enables educators to optimize instructional designs under NEP 2020, reducing superfluous load while increasing relevant processing for various learners, including those with special needs; future study should look into longitudinal applications and cross-cultural adaptations.

References

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Chen, O., Kalyuga, S., & Sweller, J. (2018). The expertise reversal effect is a variant of the more general element interactivity effect. *Educational Psychology Review, 30*(3), 771–797. https://doi.org/10.1007/s10648-017-9414-7

DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Sage.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX: Results of empirical and theoretical research. *Advances in Man-Machine Systems Research, 52*, 139–183.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2016). Cognitive load measurement as a means to advance cognitive load theory. In *Cognitive load theory* (pp. 63-71). Routledge.

Sweller, J. (1988). Cognitive load during problem solving. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and instruction*, *8*(4), 351-362.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.